A low-latency full-duplex audio over IP streamer

Asbjørn Sæbø and U. Peter Svensson
Centre for Quantifiable Quality of Service in Communication Systems
NNRS 2005
What is LDAS?

— A Low Delay Audio Streamer in software
  • basically an audio to UDP/IP adapter
  • developed for Linux, using the ALSA sound system

— A research tool
  • distributed multimedia interaction
  • perceived quality of service

— Aimed at demanding applications
  • low latency, high quality multichannel audio
Background and motivation

— Q2S: Audio over IP Networks research area
— Quality beyond voice over IP:
  • Lower latency
  • Multiple channels $\Rightarrow$ higher bit rates
  • Higher audio quality $\Rightarrow$ higher bit rates
— Existing solutions
  • No access to source code, and/or
  • Not fulfilling requirements
— Goal: Fully open software, suitable as research tool
Worldview

Quality of service, as perceived from the endpoints

Scenario: Geographically distributed musicians performing together, with audio (and video) being distributed between all the participants via a connecting IP network.
Previous related work at Q2S/NTNU

Ola Strand:
— Distributed musical performance
— Transmission of audio and video

Håkon Liestøl Winge:
— Music playing over IP
— Measured influence of latency

Otto Wittner / Sigurd Saue:
— Low-delay Windows Streamer

Trond Iver Røste Pedersen:
— Streaming in Java
Requirements

Application: Networked ensemble playing, audio for advanced video conference

— Low latency: Less than 20 - 30 ms for ensemble playing
— Multiple channels: For realistic transmission of acoustic environments and multi-channel 3D audio
— High quality: From CD-quality and upwards
Specification (minimum)

**Audio:** 44.1kHz/48kHz, 16bit, two-channel PCM, possibility for coding.

**Network:** UDP/IP, with precautions against UDP shortcomings. No retransmission.

**Latency:** Less than 20 ms (analog to analog) over LAN.

**Synchronisation:**
- Receiver and sender must be kept in sync.
- Buffering, to minimise the effects of network transmission time jitter
Implementation

Three threads: Recorder/sender, receiver and playback

- Read audio data from sound card
- Build packet, send to network
- Process audio
- N+1

- Receive packet from network
- Enqueue packet
- Build audio period from enqueued data
- N+2

- Write audio data to sound card
- N+3

www.ntnu.no
A. Sæbø and U. P. Svensson, Low-latency audio over IP
Latency

Audio delivered in *periods*

— Sender sound card buffer: One period of latency
— Receiver sound card buffer: At least one period
— Receiver queue: Adjustable
— Network transmission time
— A/D and D/A: 2 – 3 ms
— Processing latency (software, OS)
Packet format

The payload of the UDP packet:

<table>
<thead>
<tr>
<th>Audio data (one period)</th>
<th>Seq. Num.</th>
<th>Time stamp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame 1</td>
<td>Frame 2</td>
<td>Frame M</td>
</tr>
<tr>
<td>Ch. 1 sample</td>
<td>Ch. 2 sample</td>
<td>Ch. N sample</td>
</tr>
</tbody>
</table>
Packet stream control

Protocol implicitly defined by packet format, sequence numbering and receiver enqueueing combination. Handles UDP shortcomings and network transmission problems.

— Order packets
— Detect lost/missing packets
— Reject duplicates
— Reject late packets

Missing packets replaced by dummy data. (Possibility: Error concealment.)
Synchronisation

— “Large scale” synchronisation
  • Queue is a sliding window onto packet stream
  • “Early” packet (outside window) ⇒ resynchronisation
  • “Too many” late packets ⇒ resynchronisation

— Drift adjustment by watermark algorithm
  • Queue length too high ⇒ skip single samples
  • Queue length too low ⇒ reuse single samples
Results

— Multichannel transmission
  • Two channel
  • Four channel (one-way, early version)

— Full duplex transmission
  • Full duplex version
  • Double set one-way version

— Latency
  • Less than 30ms (analog to analog)
  • Less than 10ms expected for other sound interfaces
Plans and possibilities

Architecture: From sender/receiver and pairwise peer-to-peer to general architecture capable of one-way, pairwise peer-to-peer, multi-way peer-to-peer, client/server and relay solutions

Network timing: Time stamps may be used for transmission time jitter monitoring

Experiments:
- Networked ensemble playing
- Effects of latency
- Transmission of 3D audio

Coding?: Adding packet-based audio coding

Error concealment?: Instead of silence dummy data
Availability

— Open Source, GNU General Public License
— Download: http://www.q2s.ntnu.no/~asbjs/ldas/ldas.html
— Mailing list: https://pat.q2s.ntnu.no/mailman/listinfo/ldas-dev